Hand Gesture Classification

Video Labeling of Gesture-Based Biosignals

ELECOMP Capstone Design Project 2021-2022

Sponsoring Company:

Pison Technology
179 South Street 4th Floor Suite,
Boston MA 02111
https://pisontechnology.com

Company Overview:

Pison is creating the next generation of neural interfaces that focus on gesture recognition and neurophysiological insights. Pison leverages our patented electroneurography (ENG) platform to translate data generated from our body's natural physiological electricity stemming at the brain into machine interpretable events in software. This allows for never-before-seen intuitive control of any experience in the IoT universe and discovery of neurological correlations drawn from our proprietary AI and data sets. Pison's government, Department of Defense and Fortune 100 customers have a deep pipeline of pending projects fueling our growth. Pison's long term vision is to own the platform which translates human intention to machine-readable code.

Technical Directors:

Xiaofeng Tan
Director of Algorithms
xiaofeng@pison.com
https://www.linkedin.com/in/xiaofeng-tan-0a832369

Matthew Fleury
Machine Learning Team Lead
matthew@pison.com
https://www.linkedin.com/in/matt-fleury

Project Motivation:

When collecting gesture-related biosignals for researching, developing, and training gesture classification models, one of the most important preprocessing steps is identifying which regions of the signal correspond to the intentional performance of the gesture (known as "onset detection" or "activity detection"). Onset detection is used to ensure accurate labeling of the samples used to train these models: in particular, it is important to capture as much of the active signal as possible, and even more important to avoid labeling "inactive" or "resting" samples, surrounding the true active region, as "active". If this labeling is not performed, overfitting usually occurs, and the generalization ability of the model is diminished.

Purely signal-based onset detection methods exist, but are often susceptible to noise such as electromagnetic interference (EMI), and they generally perform poorly with low signal-to-noise ratio (SNR) signals. Further, these methods provide no information about which gesture was performed, so this information needs to be known in advance. Thus, we are leveraging computer vision (CV) methods to perform activity detection of biosignals based on recorded videos of the user's hand while performing the gesture. The long-term vision is to gamify this data collection experience: having the user freely perform various gestures in a Snapchat-style app, activating different filters with different gestures, etc., while simultaneously recording video and capturing device data, and thus automatically labeling the data. This would drastically lower the barrier to data collection and would thus enable Pison to amass larger amounts of data to perform future algorithm R/D and model pre-training.

Anticipated Best Outcome:

The anticipated best outcome of this project is the development of a fully automated labeling algorithm by April 2022. Note that this project is funded by a grant from the National Science Foundation (NSF). We are aiming to achieve robust accuracy (99+% overall classification accuracy across all users in the test set), but in terms of the contractual deliverable this is not strictly necessary.

Project Details:

Overall System Concept:

The first method we are developing consists of a semi-supervised algorithm that knows a priori which gesture is performed in a given collection window and is determining the exact onsets/offsets of the gesture in the signal. The second, more difficult and challenging, method we are developing involves a fully automated labeling algorithm, where the user freely performs gestures and the algorithm both determines onset/offset and assigns the gesture label. This project will focus on developing the fully automated labeling algorithm.

A video collection Android app has been developed at Pison which records video through the phone camera while capturing and saving Pison device data streaming over BLE to the phone. Following collection, this app uploads the data to various storage points on cloud services: the device data is stored in a Postgres database on Microsoft Azure, and the video data is stored in blob storage on Google Cloud Platform (GCP). This data is then read into notebook servers or Kubernetes containers on GCP to perform intermediate processing (such as filtering the device data and synchronizing the two data streams), feature extraction, and algorithm development and testing. The main goal of this project is to create this algorithm!

Block Diagram:

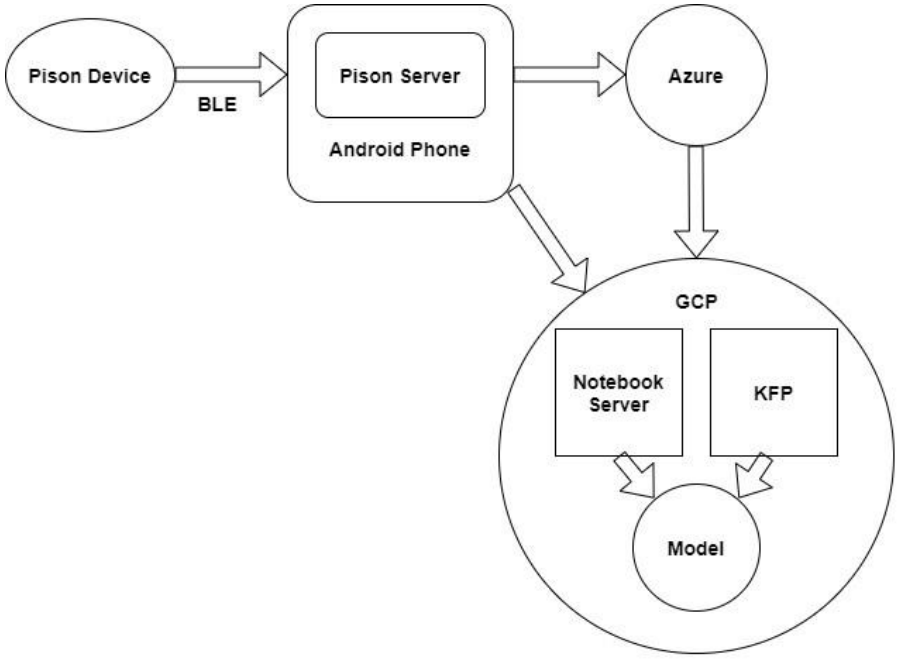

Figure 1: Data and model development/deployment infrastructure

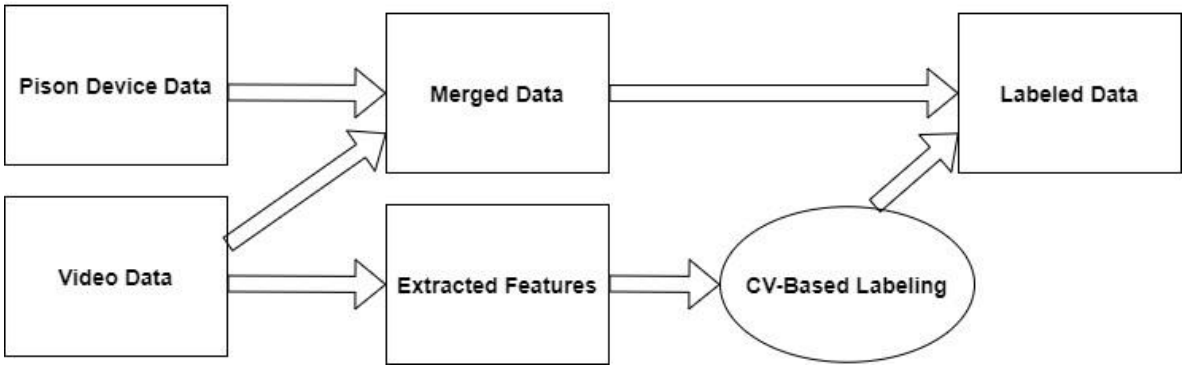

Figure 2: Model pipeline

Software/Computer Tasks:

- Collect additional video labeling app data as needed
- Validate semi-supervised labeling algorithm on any new data
- Research and develop a fully automated labeling algorithm
- Benchmark the fully automated labeling algorithm on test data to estimate its generalization performance

Composition of Team:

2 Computer Engineers, preferably with an interest in machine learning and biosignals.

Skills Required:

Computer Engineering Skills Required:

- Background in Python programming
- Familiarity with working in Jupyter notebooks
- Previous experience with using GCP is beneficial
- Familiarity with Kubernetes and Kubeflow Pipelines is a plus
- Familiarity with convolutional neural networks (CNN's) and conventional ML algorithms (such as random forests, linear discriminant analysis, etc.) is helpful
- Knowledge of the basic mathematical and statistical concepts underlying ML (linear algebra, multivariable calculus, statistical distributions, statistical tests)
- Innovative, frontiering mindset; willing to devise and try new things not necessarily found in the literature
- Ability to interface with other teams
- Positive, constructive attitude; open to critical review exercises as a team

Anticipated Best Outcome's Economic Impact on Company's Business:

The best outcome of this project would provide Pison with a first fully automated labeling algorithm to be deployed to future data collection apps. The larger amounts of data streaming in from such low-barrier-to-collection apps would magnify the amount of research, development, and product-market exploration that Pison could perform, thus enabling the company to not only make its technology more robust for existing use cases, but also rapidly expand into new use cases as product-market-fits are identified. This would enable Pison to generate more self-sufficient streams of revenue, and bolster existing ones, allowing the continued growth of the company.

Broader Implications of the Best Outcome on the Company's Industry:

Such a method of capturing gesture data may help transform the space from the current within-subject, within-session calibration paradigm into a realm of "big data", where data from thousands of users are leveraged to build powerful transfer learning algorithms that can generalize to new users and new user interfaces. This would make the technology a robust, out-of-box experience suitable for scaled commercial use in future applications.